

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 101042526A  
Source: OIP E  
Date Processed by STIC: 10/21/04

# ***ENTERED***



OIPE

## RAW SEQUENCE LISTING

DATE: 10/21/2004

PATENT APPLICATION: US/10/042,526A

TIME: 11:28:34

Input Set : A:\38150.txt

Output Set: N:\CRF4\10212004\J042526A.raw

3 <110> APPLICANT: GISSMANN, et al.

5 <120> TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE FORMULATIONS AND METHODS  
OF USE

7 <130> FILE REFERENCE: 27013/38150

9 <140> CURRENT APPLICATION NUMBER: US 10/042,526A

C--> 10 <141> CURRENT FILING DATE: 2002-04-29

12 <150> PRIOR APPLICATION NUMBER: US 09/632,286

13 <151> PRIOR FILING DATE: 2000-08-03

15 <150> PRIOR APPLICATION NUMBER: US 08/944,368

16 <151> PRIOR FILING DATE: 1997-10-06

18 <160> NUMBER OF SEQ ID NOS: 28

20 <170> SOFTWARE: PatentIn version 3.3

22 <210> SEQ ID NO: 1

23 <211> LENGTH: 1518

24 <212> TYPE: DNA

25 <213> ORGANISM: Human Papilloma Virus

28 <220> FEATURE:

29 <221> NAME/KEY: CDS

30 <222> LOCATION: (1)..(1518)

32 <400> SEQUENCE: 1

33	atg tct ctt tgg ctg cct agt gag gcc act gtc tac ttg cct cct gtc	48
34	Met Ser Leu Trp Leu Pro Ser Glu Ala Thr Val Tyr Leu Pro Pro Val	
35	1 5 10 15	
37	cca gta tct aag gtt gta agc acg gat gaa tat gtt gca cgc aca aac	96
38	Pro Val Ser Lys Val Val Ser Thr Asp Glu Tyr Val Ala Arg Thr Asn	
39	20 25 30	
41	ata tat tat cat gca gga aca tcc aga cta ctt gca gtt gga cat ccc	144
42	Ile Tyr Tyr His Ala Gly Thr Ser Arg Leu Leu Ala Val Gly His Pro	
43	35 40 45	
45	tat ttt cct att aaa aaa cct aac aat aac aaa ata tta gtt cct aaa	192
46	Tyr Phe Pro Ile Lys Lys Pro Asn Asn Asn Lys Ile Leu Val Pro Lys	
47	50 55 60	
49	gta tca gga tta caa tac agg gta ttt aga ata cat tta cct gac ccc	240
50	Val Ser Gly Leu Gln Tyr Arg Val Phe Arg Ile His Leu Pro Asp Pro	
51	65 70 75 80	
53	aat aag ttt ggt ttt cct gac acc tca ttt tat aat cca gat aca cag	288
54	Asn Lys Phe Gly Phe Pro Asp Thr Ser Phe Tyr Asn Pro Asp Thr Gln	
55	85 90 95	
57	cgg ctg gtt tgg gcc tgt gta ggt gtt gag gta ggt cgt ggt cag cca	336
58	Arg Leu Val Trp Ala Cys Val Gly Val Glu Val Gly Arg Gly Gln Pro	
59	100 105 110	
61	tta ggt gtg ggc att agt ggc cat cct tta tta aat aaa ttg gat gac	384
62	Leu Gly Val Gly Ile Ser Gly His Pro Leu Leu Asn Lys Leu Asp Asp	
63	115 120 125	

## RAW SEQUENCE LISTING

DATE: 10/21/2004

PATENT APPLICATION: US/10/042,526A

TIME: 11:28:34

Input Set : A:\38150.txt

Output Set: N:\CRF4\10212004\J042526A.raw

65	aca gaa aat gct agt gct tat gca gca aat gca ggt gtg gat aat aga	432
66	Thr Glu Asn Ala Ser Ala Tyr Ala Ala Asn Ala Gly Val Asp Asn Arg	
67	130 135 140	
69	gaa tgt ata tct atg gat tac aaa caa aca caa ttg tgt tta att ggt	480
70	Glu Cys Ile Ser Met Asp Tyr Lys Gln Thr Gln Leu Cys Leu Ile Gly	
71	145 150 155 160	
73	tgc aaa cca cct ata ggg gaa cac tgg ggc aaa gga tcc cca tgt acc	528
74	Cys Lys Pro Pro Ile Gly Glu His Trp Gly Lys Gly Ser Pro Cys Thr	
75	165 170 175	
77	aat gtt gca gta aat cca ggt gat tgt cca cca tta gag tta ata aac	576
78	Asn Val Ala Val Asn Pro Gly Asp Cys Pro Pro Leu Glu Leu Ile Asn	
79	180 185 190	
81	aca gtt att cag gat ggt gat atg gtt gat act ggc ttt ggt gct atg	624
82	Thr Val Ile Gln Asp Gly Asp Met Val Asp Thr Gly Phe Gly Ala Met	
83	195 200 205	
85	gac ttt act aca tta cag gct aac aaa agt gaa gtt cca ctg gat att	672
86	Asp Phe Thr Thr Leu Gln Ala Asn Lys Ser Glu Val Pro Leu Asp Ile	
87	210 215 220	
89	tgt aca tct att tgc aaa tat cca gat tat att aaa atg gtg tca gaa	720
90	Cys Thr Ser Ile Cys Lys Tyr Pro Asp Tyr Ile Lys Met Val Ser Glu	
91	225 230 235 240	
93	cca tat ggc gac agc tta ttt ttt tat tta cga agg gaa caa atg ttt	768
94	Pro Tyr Gly Asp Ser Leu Phe Phe Tyr Leu Arg Arg Glu Gln Met Phe	
95	245 250 255	
97	gtt aga cat tta ttt aat agg gct ggt gct gtt ggt gaa aat gta cca	816
98	Val Arg His Leu Phe Asn Arg Ala Gly Ala Val Gly Glu Asn Val Pro	
99	260 265 270	
101	gac gat tta tac att aaa ggc tct ggg tct act gca aat tta gcc agt	864
102	Asp Asp Leu Tyr Ile Lys Gly Ser Gly Ser Thr Ala Asn Leu Ala Ser	
103	275 280 285	
105	tca aat tat ttt cct aca cct agt ggt tct atg gtt acc tct gat gcc	912
106	Ser Asn Tyr Phe Pro Thr Pro Ser Gly Ser Met Val Thr Ser Asp Ala	
107	290 295 300	
109	caa ata ttc aat aaa cct tat tgg tta caa cga gca cag ggc cac aat	960
110	Gln Ile Phe Asn Lys Pro Tyr Trp Leu Gln Arg Ala Gln Gly His Asn	
111	305 310 315 320	
113	aat ggc att tgt tgg ggt aac caa cta ttt gtt act gtt gtt gat act	1008
114	Asn Gly Ile Cys Trp Gly Asn Gln Leu Phe Val Thr Val Val Asp Thr	
115	325 330 335	
117	aca cgc agt aca aat atg tca tta tgt gct gcc ata tct act tca gaa	1056
118	Thr Arg Ser Thr Asn Met Ser Leu Cys Ala Ala Ile Ser Thr Ser Glu	
119	340 345 350	
121	act aca tat aaa aat act aac ttt aag gag tac cta cga cat ggg gag	1104
122	Thr Thr Tyr Lys Asn Thr Asn Phe Lys Glu Tyr Leu Arg His Gly Glu	
123	355 360 365	
125	gaa tat gat tta cag ttt att ttt caa ctg tgc aaa ata acc tta act	1152
126	Glu Tyr Asp Leu Gln Phe Ile Phe Gln Leu Cys Lys Ile Thr Leu Thr	
127	370 375 380	
129	gca gac gtt atg aca tac ata cat tct atg aat tcc act att ttg gag	1200

## RAW SEQUENCE LISTING

DATE: 10/21/2004

PATENT APPLICATION: US/10/042,526A

TIME: 11:28:34

Input Set : A:\38150.txt

Output Set: N:\CRF4\10212004\J042526A.raw

```

130 Ala Asp Val Met Thr Tyr Ile His Ser Met Asn Ser Thr Ile Leu Glu
131 385                      390                      395                      400
133 gac tgg aat ttt ggt cta caa cct ccc cca gga ggc aca cta gaa gat      1248
134 Asp Trp Asn Phe Gly Leu Gln Pro Pro Pro Gly Gly Thr Leu Glu Asp
135                      405                      410                      415
137 act tat agg ttt gta acc tcc cag gca att gct tgt caa aaa cat aca      1296
138 Thr Tyr Arg Phe Val Thr Ser Gln Ala Ile Ala Cys Gln Lys His Thr
139                      420                      425                      430
141 cct cca gca cct aaa gaa gat ccc ctt aaa aaa tac act ttt tgg gaa      1344
142 Pro Pro Ala Pro Lys Glu Asp Pro Leu Lys Lys Tyr Thr Phe Trp Glu
143                      435                      440                      445
145 gta aat tta aag gaa aag ttt tct gca gac cta gat cag ttt cct tta      1392
146 Val Asn Leu Lys Glu Lys Phe Ser Ala Asp Leu Asp Gln Phe Pro Leu
147                      450                      455                      460
149 gga cgc aaa ttt tta cta caa gca gga ttg aag gcc aaa cca aaa ttt      1440
150 Gly Arg Lys Phe Leu Leu Gln Ala Gly Leu Lys Ala Lys Pro Lys Phe
151 465                      470                      475                      480
153 aca tta gga aaa cga aaa gct aca ccc acc acc tca tct acc tct aca      1488
154 Thr Leu Gly Lys Arg Lys Ala Thr Pro Thr Thr Ser Ser Thr Ser Thr
155                      485                      490                      495
157 act gct aaa cgc aaa aaa cgt aag ctg taa      1518
158 Thr Ala Lys Arg Lys Lys Arg Lys Leu
159                      500                      505
162 <210> SEQ ID NO: 2
163 <211> LENGTH: 505
164 <212> TYPE: PRT
165 <213> ORGANISM: Human Papilloma Virus
167 <400> SEQUENCE: 2
169 Met Ser Leu Trp Leu Pro Ser Glu Ala Thr Val Tyr Leu Pro Pro Val
170 1                      5                      10                      15
173 Pro Val Ser Lys Val Val Ser Thr Asp Glu Tyr Val Ala Arg Thr Asn
174                      20                      25                      30
177 Ile Tyr Tyr His Ala Gly Thr Ser Arg Leu Leu Ala Val Gly His Pro
178                      35                      40                      45
181 Tyr Phe Pro Ile Lys Lys Pro Asn Asn Asn Lys Ile Leu Val Pro Lys
182                      50                      55                      60
185 Val Ser Gly Leu Gln Tyr Arg Val Phe Arg Ile His Leu Pro Asp Pro
186 65                      70                      75                      80
189 Asn Lys Phe Gly Phe Pro Asp Thr Ser Phe Tyr Asn Pro Asp Thr Gln
190                      85                      90                      95
193 Arg Leu Val Trp Ala Cys Val Gly Val Glu Val Gly Arg Gly Gln Pro
194                      100                     105                     110
197 Leu Gly Val Gly Ile Ser Gly His Pro Leu Leu Asn Lys Leu Asp Asp
198                      115                     120                     125
201 Thr Glu Asn Ala Ser Ala Tyr Ala Ala Asn Ala Gly Val Asp Asn Arg
202                      130                     135                     140
205 Glu Cys Ile Ser Met Asp Tyr Lys Gln Thr Gln Leu Cys Leu Ile Gly
206 145                     150                     155                     160
209 Cys Lys Pro Pro Ile Gly Glu His Trp Gly Lys Gly Ser Pro Cys Thr

```

## RAW SEQUENCE LISTING

DATE: 10/21/2004

PATENT APPLICATION: US/10/042,526A

TIME: 11:28:34

Input Set : A:\38150.txt

Output Set: N:\CRF4\10212004\J042526A.raw

```

210          165          170          175
213 Asn Val Ala Val Asn Pro Gly Asp Cys Pro Pro Leu Glu Leu Ile Asn
214          180          185          190
217 Thr Val Ile Gln Asp Gly Asp Met Val Asp Thr Gly Phe Gly Ala Met
218          195          200          205
221 Asp Phe Thr Thr Leu Gln Ala Asn Lys Ser Glu Val Pro Leu Asp Ile
222          210          215          220
225 Cys Thr Ser Ile Cys Lys Tyr Pro Asp Tyr Ile Lys Met Val Ser Glu
226 225          230          235          240
229 Pro Tyr Gly Asp Ser Leu Phe Phe Tyr Leu Arg Arg Glu Gln Met Phe
230          245          250          255
233 Val Arg His Leu Phe Asn Arg Ala Gly Ala Val Gly Glu Asn Val Pro
234          260          265          270
237 Asp Asp Leu Tyr Ile Lys Gly Ser Gly Ser Thr Ala Asn Leu Ala Ser
238          275          280          285
241 Ser Asn Tyr Phe Pro Thr Pro Ser Gly Ser Met Val Thr Ser Asp Ala
242          290          295          300
245 Gln Ile Phe Asn Lys Pro Tyr Trp Leu Gln Arg Ala Gln Gly His Asn
246 305          310          315          320
249 Asn Gly Ile Cys Trp Gly Asn Gln Leu Phe Val Thr Val Val Asp Thr
250          325          330          335
253 Thr Arg Ser Thr Asn Met Ser Leu Cys Ala Ala Ile Ser Thr Ser Glu
254          340          345          350
257 Thr Thr Tyr Lys Asn Thr Asn Phe Lys Glu Tyr Leu Arg His Gly Glu
258          355          360          365
261 Glu Tyr Asp Leu Gln Phe Ile Phe Gln Leu Cys Lys Ile Thr Leu Thr
262          370          375          380
265 Ala Asp Val Met Thr Tyr Ile His Ser Met Asn Ser Thr Ile Leu Glu
266 385          390          395          400
269 Asp Trp Asn Phe Gly Leu Gln Pro Pro Pro Gly Gly Thr Leu Glu Asp
270          405          410          415
273 Thr Tyr Arg Phe Val Thr Ser Gln Ala Ile Ala Cys Gln Lys His Thr
274          420          425          430
277 Pro Pro Ala Pro Lys Glu Asp Pro Leu Lys Lys Tyr Thr Phe Trp Glu
278          435          440          445
281 Val Asn Leu Lys Glu Lys Phe Ser Ala Asp Leu Asp Gln Phe Pro Leu
282          450          455          460
285 Gly Arg Lys Phe Leu Leu Gln Ala Gly Leu Lys Ala Lys Pro Lys Phe
286 465          470          475          480
289 Thr Leu Gly Lys Arg Lys Ala Thr Pro Thr Thr Ser Ser Thr Ser Thr
290          485          490          495
293 Thr Ala Lys Arg Lys Lys Arg Lys Leu
294          500          505
297 <210> SEQ ID NO: 3
298 <211> LENGTH: 297
299 <212> TYPE: DNA
300 <213> ORGANISM: Human Papilloma Virus
303 <220> FEATURE:
304 <221> NAME/KEY: CDS

```

## RAW SEQUENCE LISTING

DATE: 10/21/2004

PATENT APPLICATION: US/10/042,526A

TIME: 11:28:34

Input Set : A:\38150.txt

Output Set: N:\CRF4\10212004\J042526A.raw

```

305 <222> LOCATION: (1)..(297)
307 <400> SEQUENCE: 3
308 atg cat gga gat aca cct aca ttg cat gaa tat atg tta gat ttg caa      48
309 Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln
310 1          5          10          15
312 cca gag aca act gat ctc tac tgt tat gag caa tta aat gac agc tca      96
313 Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser
314          20          25          30
316 gag gag gag gat gaa ata gat ggt cca gct gga caa gca gaa ccg gac      144
317 Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro Asp
318          35          40          45
320 aga gcc cat tac aat att gta acc ttt tgt tgc aag tgt gac tct acg      192
321 Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser Thr
322          50          55          60
324 ctt cgg ttg tgc gta caa agc aca cac gta gac att cgt act ttg gaa      240
325 Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu
326 65          70          75          80
328 gac ctg tta atg ggc aca cta gga att gtg tgc ccc atc tgt tct cag      288
329 Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln
330          85          90          95
332 aaa cca taa      297
333 Lys Pro
337 <210> SEQ ID NO: 4
338 <211> LENGTH: 98
339 <212> TYPE: PRT
340 <213> ORGANISM: Human Papilloma Virus
342 <400> SEQUENCE: 4
344 Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln
345 1          5          10          15
348 Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser
349          20          25          30
352 Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro Asp
353          35          40          45
356 Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser Thr
357          50          55          60
360 Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu
361 65          70          75          80
364 Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln
365          85          90          95
368 Lys Pro
372 <210> SEQ ID NO: 5
373 <211> LENGTH: 34
374 <212> TYPE: DNA
375 <213> ORGANISM: Artificial sequence
377 <220> FEATURE:
378 <223> OTHER INFORMATION: Synthetic primer
380 <400> SEQUENCE: 5
381 ccccgatatac gcctttaatg tataaatcgt ctgg      34
384 <210> SEQ ID NO: 6

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/042,526A

DATE: 10/21/2004

TIME: 11:28:35

Input Set : A:\38150.txt

Output Set: N:\CRF4\10212004\J042526A.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date